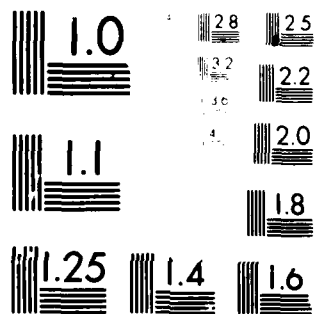


AD-A084 244

ARIZONA STATE UNIV TEMPE COLL OF EDUCATION F/G 5/9
THE EFFECTS OF SELF-EVALUATION AS AN INSTRUCTIONAL FEEDBACK STR--ETC(U)
DEC 78 L M CAREY, L ISRAELITE, R F SCHMID AFOSR-76-2900
UNCLASSIFIED TR-81230 AFOSR -TR-80-0274 NL

[OR]
AD
AD-84 244

END
DATE
FILMED
6-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

EDSR-TR- 80 - 0274

CENTER FOR INSTRUCTIONAL AND
TRAINING TECHNOLOGY

LEVEL II

Handwritten signature/initials

THE EFFECTS OF SELF-EVALUATION AS
AN INSTRUCTIONAL FEEDBACK STRATEGY

Lou M. Carey
Larry Israelite
Richard F. Schmid

Approved for public release;
distribution unlimited.

A



USAF Office of Scientific Research

Grant No. 76-2900

ARIZONA STATE UNIVERSITY
EDUCATIONAL TECHNOLOGY

Technical Report #81230

80 5 14 094

ADA084244

FILE COPY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) AFOSR-TR-80-0274	2. GOVT ACCESSION NO. AD-A084244	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) The effects of self-evaluation as an instructional feedback strategy.		5. TYPE OF REPORT & PERIOD COVERED INTERIM
7. AUTHOR(s) (10) Lou M. Carey, Larry Israelite, Richard F. Schmid		6. PERFORMING ORG. REPORT NUMBER 81230
9. PERFORMING ORGANIZATION NAME AND ADDRESS Arizona State University Tempe, AZ 85281		8. CONTRACT OR GRANT NUMBER(s) AFOSR-76-2900
11. CONTROLLING OFFICE NAME AND ADDRESS AF Office of Scientific Research USAF Bolling AFB, DC		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F (16) 2313 (17) A2
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (15) AFOSR-76-2900		12. REPORT DATE 30 Dec 78
		13. NUMBER OF PAGES 28
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. (12) 34		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) (9) Technical Repts		
18. SUPPLEMENTARY NOTES (14) TR-7423A		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Feedback Instructional systems design Self-evaluation Instructor feedback Instruction		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The purpose of this study was to investigate the effects of (a) student self-evaluation and (b) instructor feedback on the self-evaluations on the ability of students to evaluate their own materials and on their performance. Fifty-six education students were rated with respect to materials which they produced in an instructional media course. Students in the first experimental group used a check list to evaluate their own products prior to submitting them for grading by the instructor. Students in the second		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

409989

UNCLASSIFIED

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

experimental group did the same, but received comments from the instructor regarding their self-evaluation. Students in the control group did not engage in any form of self-evaluation.

The relationship between student score predictions and instructor scores on the final class product was reasonably high for the control group. However, the combination of self-evaluation and instructor feedback on self-evaluation raised this figure to .81. Criteria specific feedback on self-evaluation seems to be an effective way to develop student skills in estimating the quality of their performance.

It was anticipated that self-evaluation itself without any instructor feedback on the accuracy of the evaluation would raise the correlation between student and instructor scores. The analysis revealed that self-evaluation itself had no beneficial effects, and in fact, seemed to have a negative effect. It is our contention that self-evaluation (i.e., the effective use of objectives) is contingent upon an ongoing process of critical comparison.

When tied to student confidence in their materials, there was a significant difference between group means. When students received instructor feedback on the quality of their materials and on their own evaluations, they had the opportunity to learn not only what instructors thought of their materials, but how their self-evaluations compared to instructor evaluations on an objective-by-objective basis.

When student confidence in materials was correctly low, feedback on self-evaluation seemed to have a negative effect on performance. The results of this study indicate that when students were unsure of their work, additional feedback only caused more confusion.

Research is needed in order to clarify the effects of feedback in learning situations in which student performances are complex and student confidence is difficult to gauge.

Further research is indicated to determine the effect of further explanation regarding the negative effect of self-evaluation only on students.

A second question for further investigation is where to begin instruction for students who are unable to use prespecified criteria to improve the quality of their work.

If trainees can learn to use prespecified objectives and criteria effectively to evaluate and shape their performances, the case for the use of instructional objectives is made even stronger.

If trainee self-evaluation can be developed to an acceptable level, then time spent on certain types of instructor feedback can be greatly reduced, freeing trainers for other instructional tasks during the training process.

Rule Learning and Systematic Instruction in
Undergraduate Pilot Training

Vernon S. Gerlach, Principal Investigator

THE EFFECTS OF SELF-EVALUATION AS
AN INSTRUCTIONAL FEEDBACK STRATEGY

Lou M. Carey
Larry Israelite
Richard F. Schmid

AIR F

R

T

1

A

1

Technical Report #81230

Research sponsored by the Air Force Office of Scientific Research,
Air Force Systems Command, USAF, under Grant No. AFOSR 75-2900. The
United States Government is authorized to reproduce and distribute
reprints for governmental purposes notwithstanding any copyright
notation hereon.

College of Education ✓
Arizona State University
Tempe, Arizona

December, 1978

Accession for	
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

A

**THE EFFECTS OF SELF-EVALUATION AS
AN INSTRUCTIONAL FEEDBACK STRATEGY**

The Effects of Self-Evaluation as an Instructional Feedback Strategy

In an effort to maximize cost effectiveness of training programs, all the military services currently use an "objectives based" instructional approach. Trainees are told what performance is expected of them and the criteria by which their performance will be judged. It is assumed they will use the criteria to evaluate the quality of their own performance and alter the performance based on corrective feedback.

When job performance, rather than class performance, is at issue, the ability to self-evaluate takes on increased importance. If individuals are unable to evaluate their own job performance or to discriminate between a correct and an incorrect performance, the results of training will be less than optimal. The question to be considered is whether trainees are able to use objectives and criteria to evaluate their own performance, and if so, whether the ability to self-evaluate improves with practice.

Experimental situations were created to investigate the effects of two self-evaluation procedures: (a) learners' evaluation of their own products using specified criteria and (b) learners' self-evaluation with instructor feedback on the accuracy of their evaluations. Two dependent variables were used to measure these effects: (a) learners' ability to use stated criteria for judging their own products and (b) learners' performance on course objectives.

Review of Literature

Literature on systematically designed instruction and the variety of instructional strategies that can be employed to improve learning is

abundant. For the purposes of this study, we narrowed our focus to research dealing with (a) systematically designed instructional procedures related to objectives, standards, and practice and feedback and (b) self-evaluation as an instructional strategy.

Systematic Instruction. Literature was surveyed on systematic instructional design procedures and the effects of using systematically designed instruction. Particular areas studied were performance objectives, performance standards, and practice and feedback.

A number of learning theorists state that learning will improve as a result of pre-specified objectives (Gagné & Briggs, 1974; Gerlach & Ely, in press; Kibler, Barker & Miles, 1970; Kibler & Basset, 1977; Mager, 1961). However, there is little research which examines the development of strategies that will increase the ability of students themselves to use objectives in learning situations. One question which remains unanswered is whether students can effectively use statements of objectives to improve their own learning effectiveness and, if not, what instructional techniques preceding instruction can be employed along with objectives to facilitate learning.

The specification of performance standards is widely accepted as an integral part of an instructional objective. If performance evaluation is judgmental, i.e., there is no clearcut right or wrong performance, then the specification of performance criteria becomes appropriate. When a simple right or wrong response is not possible, Dick and Carey (1978) suggest that instructional objectives should include a checklist of the types of behaviors which are expected when the performance of students is judged. This breakdown of behavior should give students a clearer understanding of the required performance. However, the degree to which

learners are able to use criteria in shaping their behaviors has not yet been established (Carey, 1976).

The use of practice and feedback are universally accepted as important aspects of instruction. Both are regularly included in systematically designed instructional materials. Research on practice and feedback were reviewed in this study relative to self-evaluation or student evaluation of their own work. Kulhavy, Yekovich and Dyer (1976) have shown that in programmed instruction, certain conditions maximize the effects of instructor feedback on learner performance. They found that feedback is most effective when a learner believes a test response is correct when, in fact, it is incorrect. In this case, learner response confidence is incorrectly high. When a learner believes a response is incorrect, and it is, Kulhavy et al. state that feedback is not effective because learners do not understand either the subject matter being tested, the test question, or both.

In programmed instruction, feedback indicates to the learner whether a particular response to a program frame is correct or incorrect, and the effects of feedback are measured by a subsequent test covering the same or similar material. In some instructional situations, the feedback given on one performance is expected to help the learner with future performances which are similar. When complex, multi-dimensional performances are given, learners' confidence in this performance can be measured by having them evaluate each aspect of their own performance. Self-evaluation becomes an indicator of learner response confidence. If instructor feedback is then given on the quality of learner self-evaluations, in addition to the performance of tasks to be learned, a double feedback condition exists. The effects of feedback on student performance should be noticed under these conditions regardless of learner confidence, especially when the feedback

given on performance and on self-evaluations is instructional and causes students to attend more to objectives and standards for acceptable performance.

Self-evaluation. One method for determining the degree to which learners are able to use objectives and criteria statements to shape their performances is to ask learners to use them to evaluate their own work. Those who are able to accurately evaluate their work using pre-specified criteria probably possess a clear understanding of the objective and criteria concerned. On the other hand, learners who are unable to use performance criteria to judge the quality of their work after its completion are, likewise, unable to use the same criteria during its production. Learners who are unable to use criteria during production would not be expected to perform as well as those who have the ability to apply statements of criteria when developing instructional products.

Clark (1938) found that students were able to evaluate their own performance on college level algebra, quantitative analysis, and chemistry problems. The correlation between student scores and instructor scores was .80. Bennet (1958) found further support for the argument that students are able to evaluate their own performances.

Both Clark and Bennet reported that students gave themselves the same grade as their instructors in a high percentage of cases. Estimating a single grade, or single number on a 5-point scale, would seem to be an easier task than determining how closely several parts of a complex production match the specified performance criteria. When one performs in the field, evaluations are based on how closely the actual performance approximates the expected performance.

The question of whether self-evaluation actually promotes learning remains unanswered. McEwen (1957) found that learning was unaffected by students' evaluation of their materials. In this study, however, performance had no effect on course grades, so students had little at stake. Had the performance affected grades, student interest in learning might have been greater and results might well have been different.

In a study involving sculpture, graphics, painting, and drawing skills, Fried (1965) found that although sculpture improved as a result of self-evaluation, the other skills studied did not. Noting the inconclusiveness of his own finding, Fried concluded that the value of self-evaluation was still undetermined, and recommended continued research in the area.

Duel (1958), in a study conducted in two Air Force technical schools, found that achievement was improved when students were given formal and periodic opportunities to evaluate their own work. A study in which self-evaluation had a positive effect on students' mechanical drawing skills was conducted by Irwin (1973). He found that students who evaluated their own materials throughout the school year learned more as measured on a standardized test which assessed mechanical drawing skills. In this study, however, results were confounded because the self-evaluating group had a significantly higher mean IQ than the control group.

Evidence has been cited which both supports and contradicts the assertion that self-evaluation improves student performance. Little has been written on the effects of (a) student practice on evaluating their own work and of (b) instructor feedback on student self-evaluation. The combination of systematic instructional procedures with self-evaluation as an additional feedback strategy may prove to make instruction more effective.

Hypotheses

The literature concerning learner self-evaluation, though limited, does indicate that students may possess at least a limited ability to evaluate their own work and that periodic self-evaluation may improve learner performance.

The specific hypotheses tested in this research are the following:

1. The correlation between learner and instructor evaluation scores on a final course project will be higher when instructor feedback is given on previous learner self-evaluations throughout the term than when no instructor feedback on self-evaluation is given, or when students do not participate in self-evaluation throughout the term.
2. Learner performance on a final class project will be better when students evaluate their own materials throughout the term and receive instructor feedback on their self-evaluations than when students do not receive feedback on their self-evaluations, or do not participate in self-evaluation throughout the term.
3. For a subgroup of students who incorrectly assume that their initial products are well produced, learner performance on a final product will be better when feedback is given on the quality of the students' evaluation of their own work than when instructor feedback is given only on the quality of the product.
4. For a subgroup of students who correctly assume their product is not well produced, learner performance on a final product will be better when feedback is given on the quality of the initial product and on student evaluations of the project than when feedback is given only on the product.

Pilot Study

In an attempt to identify experimental procedures which could most easily be adapted to an ongoing instructional program, a pilot study was conducted during the spring semester of 1978 at Arizona State University, Tempe.

Method

Sample. Three instructors and nine undergraduate media production classes participated in the study. Three intact classes were assigned to each of the three treatment groups: control, self-evaluation and feedback. Each instructor was responsible for only one treatment group.

Procedures. The control group used performance objectives, criteria statements, and instructor feedback to produce a mediated unit of instruction. The self-evaluation group, in addition to the above, evaluated their own materials using the instructor evaluation form prior to submission. Instructor feedback on the quality of the materials was the same for both the control and the self-evaluation groups. The third or feedback group produced their materials in the same manner as the other two groups. Like the self-evaluation group, prior to submitting their materials, they evaluated their own work and submitted the evaluation forms to their instructors. However, for this group, instructor feedback consisted of both feedback on the quality of the products and on the quality of the student evaluations.

To enable comparisons of self-evaluations among the three groups, the control group used the evaluation forms to assess their final products at the end of the semester.

There were three major checkpoints during the semester at which students formally submitted their materials for instructor evaluation and

feedback: planning document evaluation, materials evaluation and mass media project evaluation. The instrument used for each evaluation is included in Appendix A.

Students were asked to complete and submit for evaluation a planning document that would describe the product they would develop during the semester. There were six areas to be evaluated on the planning document: the content outline, the purpose statement, the audience analysis, the objectives, media selection, and the use for the instruction. There were two to four questions under each of the six areas which were used as criteria to judge the merit of the planning document.

The second checkpoint was materials evaluation which occurred about two-thirds into the semester. Seven categories were to be evaluated: lettering, illustration, mounting, transparency making, construction, appearance, and other. The category marked other was used to provide evaluation flexibility to accommodate the variety of materials developed by students. Two to five questions within each category were used as criteria to assess the quality of materials relative to that category.

The third checkpoint, mass media project evaluation, was the end of the semester evaluation of the course project. The categories included were synonymous with those used in the planning document, except at this checkpoint, they were used to evaluate course products and not plans for products. The six categories each had two to five questions which were used as criteria to evaluate the product relative to that category.

Results. Pearson product moment correlations were used to compare the evaluation scores of students with those of their instructors in the control, self-evaluation, and feedback groups on the final product. Significant correlations ($p \geq .05$) were observed between student and

instructor ratings in both the control and the feedback groups. A slightly negative, but non-significant correlation was observed in the self-evaluation group.

A one-way analysis of variance among groups revealed no significant differences ($p \geq .05$) on performance scores.

Discussion

As a result of the pilot study, a number of procedural problems were identified. Due to the nature of the procedural problems, data collected during this study were judged not valid, nor generalizable to other instructional settings. Changes in procedures which were identified as being problematic are described in the following paragraphs.

The first procedure which was found to be unacceptable was the assignment of one treatment to each instructor. Training sessions were held to establish scoring guidelines, and a high degree of inter-rater reliability was established. However, during the study, which spanned an entire semester, a distinct pattern of instructor bias developed. One instructor seemed to be more lenient in grading than did the others, and also offered students more assistance in the development of their products. Students were assured of high grades by one instructor prior to formal instructor evaluation. As a result, students knew their scores before submitting their products, and the relationship between student and instructor scores was no doubt inflated. It was determined that if each instructor delivered all three treatments in a control setting, then the effects of this type of instructor bias would be minimized.

A second problem which was encountered concerned the instruments used for both student and instructor evaluations. The objectives list and grading criteria were on separate forms and although students had both

forms in their possession at the time of self-evaluation, they rarely referred to the criteria sheet when grading their products. The objectives and criteria were combined into one form to help students focus on the criteria when assessing the quality of their materials.

The pilot study helped refine experimental procedures and enabled the researchers to conduct the main study in an ongoing instructional setting. With the exception of instructor treatment assignment and evaluation instrument design, the procedures used in the pilot study were also used in the main study.

Method

Design

In this study a posttest-only control group design was used to measure the effects of student self-evaluation of instructional products, and instructor feedback on the quality of self-evaluation on student performance on the final class product. An analysis of variance at a significance level of .05 was used to measure the effects of these two independent variables. A Pearson product moment correlation at a significance level of .05 was used to measure the relationship between student evaluations and instructor evaluations of all instructional products.

Sample

The study was conducted during the 1978 summer session at Arizona State University, Tempe. Student participants were enrolled in IME 411, Audiovisual Materials and Procedures in Education. A total of 56 preservice and inservice teachers in three class sections participated.

Procedure

Students in each section were blocked according to three levels of grade point average--low, average, and high, and then randomly assigned to either the control, first, or second treatment group. At the onset of the study, there were 20 students in the control group, 16 in the first treatment group, and 20 in the second treatment group. During the conduct of the study, eight students were dropped from the control group due to procedural problems. Each course instructor taught the control and the two treatment groups in order to avoid instructor bias.

Mastery instructional strategies including use of performance objectives, specified criteria, examples of acceptable products, relevant practice, and instructor feedback on the quality of students' products were the instructional techniques used in all three treatment groups to help students produce one mediated and one non-mediated unit of instruction. The instruction, practice and feedback, and assessment procedures occurred throughout the course. There were three major checkpoints at which students formally submitted their products for instructor evaluation and feedback. The evaluation form included a listing of all aspects of the products to be evaluated and the specific criteria to be used in the evaluation. (Sample instructor evaluation forms can be seen in Appendix A.

The control group followed the instructional procedures outlined previously to produce their materials during the course. Students in the first treatment group followed these same procedures, and in addition, prior to submitting their materials for instructor evaluation at each checkpoint, students in this group used the same form as that used by instructors to evaluate their own materials. The self-evaluations were submitted along

with the materials and were not seen again by students. This treatment group is referred to as the self-evaluation group.

The second treatment group, referred to as the feedback group, followed the same procedures as the self-evaluation group with the addition of one procedure. When instructor feedback was given on the quality of student materials, students were also given feedback on their self-evaluations. Instructors asked students to describe how they graded themselves based on the criteria, and discussed how student evaluations compared with instructor evaluations.

Students in both the control and self-evaluation groups received only one type of feedback--feedback on the quality of their materials. Students in the feedback group received feedback on the quality of their materials and also on their self-evaluations. It is important to note that the instructors did not see the student self-evaluations until after student products were graded. This eliminated instructor bias caused by student perceptions of the quality of their work.

Performance Standards

Each aspect of student products was graded by course instructors on a scale of 1 to 10, with 10 being the highest score possible. Grading criteria were established by course instructors prior to the beginning of the study. Instructor scores were used as the standard against which all comparisons were made.

Results

The process of self-evaluation and the use of feedback on the accuracy of self-evaluation were studied to ascertain their effects on (a) students' evaluation of their own performance and (b) student achievement.

Student Evaluation

It was hypothesized in this study that the correlation between student evaluation scores and instructor evaluation scores on the final course project would be higher when instructor feedback on student evaluations of other class projects was given than when no feedback on student evaluations was given, or when no student evaluation took place.

Student and instructor evaluation means and standard deviations for the final course project and the correlation between them are shown in Table 1. For the control group, the student evaluation mean was 57.9 (maximum score = 60) and the instructor mean was 56.82, yielding a difference of 1.08 points with the correlation between them being .68 ($p \geq .05$). For the self-evaluation group, the student mean was 58.5, with the instructors giving a mean score of 53.67, differing by 4.83 points. The correlation between these two scores was $-.07$, which was not significant. Students in the feedback group predicted a mean score of 55.70 while the instructors' mean equaled 54.75, yielding a difference of .95 points. The correlation between these scores was $.81$ ($p \geq .05$).

Figure 1 contains an illustration of the correlations between student and instructor scores at the first and third checkpoints for the self-evaluation and feedback groups. There was very little change in student ability to evaluate their own work in the self-evaluation group. At the

Table 1

Student and Instructor Mean Performance Scores
and Standard Deviations on Final Class Project and
Correlation Between the Two Means for Each Treatment Group

Group	Student n ^a	Instructor	Standard Deviation	Correlation
Control	57.90 (11)	56.82	3.65	.68*
Self-evaluation	58.50 (14)	53.67	5.57	-.07
Feedback	55.70 (20)	54.75	8.26	.81**

Note. Maximum score = 60.

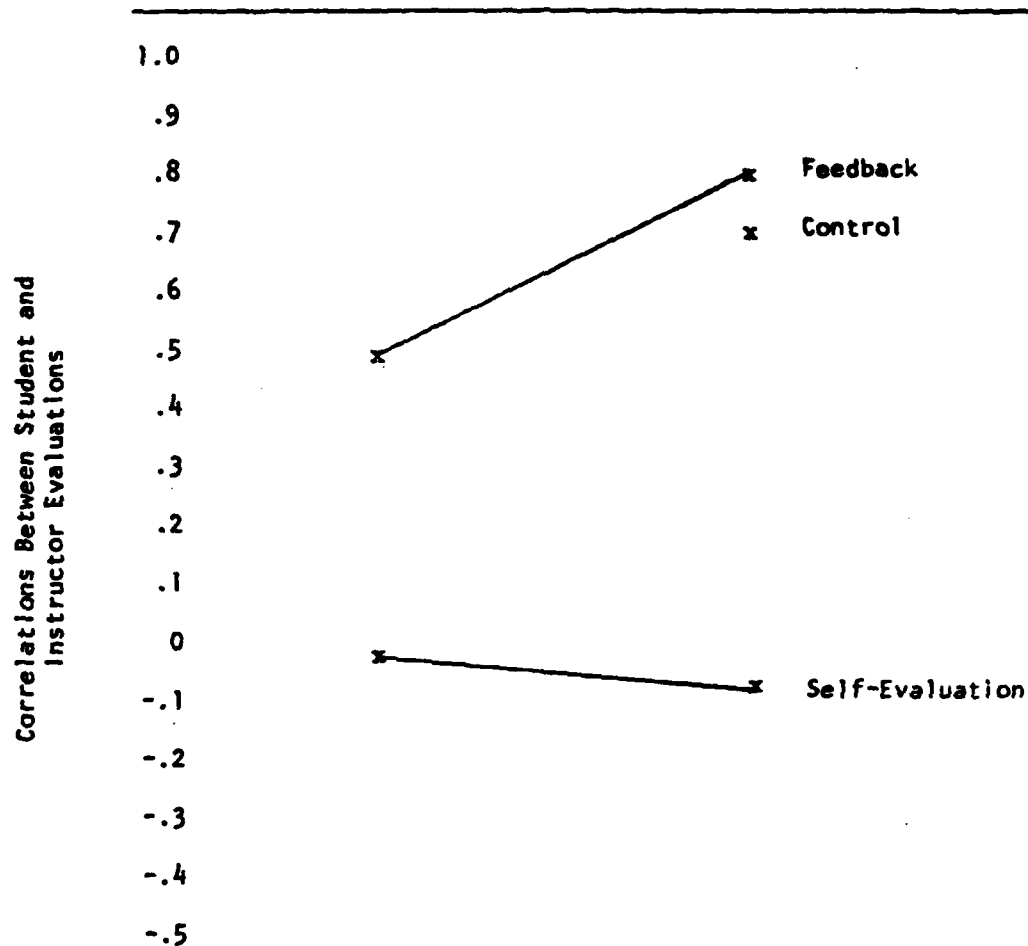
^aNumbers in parentheses indicate the number of students in each group.

* $p < .01$.

** $p < .001$.

Figure 1

Student and Instructor Evaluation Correlations
on the First and Final Evaluations



	First Evaluation	Final Evaluation
Control	--	.68
Self-evaluation	-.03	-.07
Feedback	.49	.81

first checkpoint, the correlation between student and instructor scores was $-.048$ and at the third checkpoint, $-.076$. Neither of these correlations was significant. There was little association at the outset of the study and only slightly more at the conclusion.

On the other hand, the correlation between student and instructor scores for the feedback group was $.49$ ($p < .05$) at the first evaluation checkpoint and $.81$ ($p < .05$) at the third. The relationship between the two sets of scores for this group increased approximately 65 percent.

Student Performance

It was predicted that students' performance on the final class project would be better when students evaluated previous class products and received instructor feedback on the self-evaluations than when they completed self-evaluations alone or did not complete self-evaluations at all. A one-way analysis of variance revealed no significant difference among group means.

To measure the relationship between student confidence in their materials as measured by their self-evaluation scores at the first evaluation checkpoint, and the effect of instructor feedback on student self-evaluations, students were blocked into two groups: (a) student evaluations above 50 indicating a high degree of confidence in their materials and instructor evaluations below 50 on the same materials, and (b) both student and instructor evaluations below 50 indicating low student confidence in their materials. Table 2 shows the means and standard deviations for these two subgroups for both the self-evaluation and feedback groups.

A one-way analysis of variance indicated that when student confidence was incorrectly high, there was a significant difference between the self-evaluation and feedback groups $F = 9.03$ (1, 21) $p \leq .01$. When student

Table 2
 Instructor Mean Performance Scores
 and Standard Deviations on Final Project
 Blocked on Student Confidence in Their First Project

Group	Mean Score	n ^a	Standard Deviation
Incorrect high confidence			
Self-evaluation	53.09	(10)	5.65
Feedback	58.25	(12)	1.82
Correct low confidence			
Self-evaluation	55.25	(4)	5.85
Feedback	49.50	(8)	11.29

Note. Maximum score = 60.

^aNumbers in parentheses indicate the number of students in each group.

confidence in materials was correctly low, the feedback group mean was approximately six points lower than the mean score of the self-evaluation group.

Discussion

The purpose of this study was to investigate the effects of (a) student self-evaluation and (b) instructor feedback on the self-evaluations on the ability of students to evaluate their own materials and on their performance. Although criteria statements are often included in performance objectives, there has been little evidence which indicates that students can always use criteria to shape their performance (Carey, 1976). The instruction used in this study included objectives with criteria statements and relevant practice and feedback. It conformed to the prescription for well developed instruction as specified by Gagné and Briggs (1974), Dick and Carey (1978), and Gerlach and Ely (in press).

The relationship between student score predictions and instructor scores on the final class product was reasonably high for the control group, $r = .68$ ($p \leq .05$), supporting the work of Clark (1938) and Bennet (1969). However, the combination of self-evaluation and instructor feedback on self-evaluation raised this figure to .81, indicating that the skill of self-evaluation can be improved. These findings support the hypotheses stated in this paper. Criteria specific feedback on self-evaluation seems to be an effective way to develop student skills in estimating the quality of their performance.

The correlation between student and instructor evaluation scores for the control group was much higher at the third evaluation checkpoint than

the correlation for either the self-evaluation or feedback group at the first checkpoint. For the control group the final product was their first opportunity to evaluate their own work. One possible explanation for this phenomenon is that as students go through a course, they acquire an understanding of instructors' expectations and how well their performances meet these expectations. Students then adjust their performances accordingly.

It was anticipated that self-evaluation itself without any instructor feedback on the accuracy of the evaluation would raise the correlation between student and instructor scores. The analysis revealed that self-evaluation itself had no beneficial effects, and in fact, seemed to have a negative effect. Self-evaluation seemed to confuse students and hamper their ability to use prespecified criteria to assess the quality of their materials. An explanation for this effect seems clear. Obviously, accurate evaluation must be conducted with full knowledge of both the performance and the criteria. The self-evaluation group was working at a disadvantage on both counts. The control group, as stated above, had the benefit of an entire semester's interaction with the task and instructor, and there is little doubt that college students are extremely skillful at "giving the teacher what s/he wants." The feedback group was provided the additional information as to how accurately they themselves were evaluating their work. While one might expect the self-evaluation group to perform at least as well as the control group, our hypotheses clearly suggest otherwise. It is our contention that self-evaluation (i.e., the effective use of objectives) is contingent upon an ongoing process of critical comparison. If this comparison process is disrupted, the proper use of objectives, and perhaps even overall achievement, will suffer. In the case of the self-evaluation

group, the "true" purpose of self-evaluation was suspect, and students would logically assume that their own evaluations would influence the instructor's judgment. One would therefore predict their estimates to be inordinately high. The data support this explanation, with the self-evaluation group rating themselves higher than either of the other groups. In fact, their evaluations were only two percent below the maximum; the students wishing to remain "realistic." A low correlation would be expected because the process employed by the students was at cross-purposes with that of the instructor or the research design.

The effects of self-evaluation, and self-evaluation and feedback on actual student performance is difficult to gauge in this study. Students in all three groups achieved a grade of A on the final project scoring an average of 55 out of a possible 60 points. There was little variability in student performance.

When tied to student confidence in their materials, there was a significant difference between group means. Feedback on self-evaluation of earlier material seems to be highly effective in promoting improved student scores on later performances when student confidence in materials was incorrectly high. When students received instructor feedback on the quality of their materials and on their own evaluations, they had the opportunity to learn not only what instructors thought of their materials, but how their self-evaluations compared to instructor evaluations on an objective-by-objective basis. This technique pushed the mean scores of students in the unwarranted high confidence subgroup 3.5 points above the feedback group mean from 54.75 to 58.25. This tends to support the conclusions of Kulhavy, Yekovich & Dyer (1976) that feedback is most effective when

a high degree of student confidence in materials is unwarranted. In the self-evaluation group, the difference between the mean scores of students who had incorrectly assumed their materials were good and the entire self-evaluation group was minimal, lending more support to the above argument and the inference that the addition of feedback on the self-evaluations is beneficial especially when learner confidence in their work is incorrectly high.

When student confidence in materials was correctly low, feedback on self-evaluation seemed to have a negative effect on performance, lowering the subgroup means from 54.75 to 49.50. These results contradict the hypothesis stated earlier that feedback on self-evaluations is effective for all students, but support the findings of Kulhavy et al. who concluded that when student confidence was correctly low, they were confused, and feedback was not effective. The results of this study indicate that when students were unsure of their work, additional feedback only caused more confusion.

Duel (1968) and Irwin (1973) both concluded that self-evaluation did improve student performance. Due to the ceiling effect noticed in this study, the current results can neither support their conclusions, nor refute the findings of McEowen, who found no improvement, and Fried (1965), who found improvement in some situations. It is clear that the effects of self-evaluation on student performance would possibly be more informative under conditions in which student grades are normally distributed.

Although the research seems to support the results of the Kulhavy et al. (1976) research concerning the effects of feedback, continued research in this area is needed in order to clarify the effects of feedback in learning situations in which student performances are complex and student confidence is difficult to gauge.

There were many questions raised during this study that will need to be investigated in future studies. Two were of particular interest. The first question relates to further explication regarding the negative effect of self-evaluation only on students.

A second question of interest is where to begin instruction for students who are unable to use prespecified criteria to improve the quality of their work. If students cannot perform specified tasks, cannot use objectives and criteria statements to evaluate their own work, and cannot use instructor feedback to improve their products, then the problem may be that these students do not possess prerequisite skills for the instruction. A feedback strategy that includes information on current instruction and on prerequisite skills instruction should be investigated in an attempt to identify the best strategy for this particular subgroup of students.

If trainees can learn to use prespecified objectives and criteria effectively to evaluate and shape their performances, the case for the use of instructional objectives is made even stronger. Then, not only will trainees know precisely what is expected, but they will be able to tell if and when they are able to perform the requested tasks. Learner achievement would be expected to improve, and evaluation procedures would be assisted.

If trainee self-evaluation can be developed to an acceptable level, then time spent on certain types of instructor feedback can be greatly reduced, freeing trainers for other instructional tasks during the training process.

References

- Bennett, K. B. Student evaluation--the new way. Industrial Arts and Vocational Education, 1958, 47, 300-302.
- Carey, L. M. The investigation of self-evaluation procedures for identifying the instructional needs of teachers. (Doctoral Dissertation, Florida State University, 1976). Dissertation Abstracts International, 1977, 37, 6428A. (University Microfilms No. 77-8575)
- Clark, P. E. Can college students evaluate themselves? School and Society, 1938, 47, 614-16.
- Dick, W., & Carey, L. M. The systematic design of instruction. Glenview, IL: Scott, Foresman and Company, 1978.
- Duel, H. S. Effect of periodic self-evaluation on student achievement. Journal of Educational Psychology, 1958, 49, 1979.
- Fried, H. C. Effect of end product self-evaluation on college students' art performance. (Doctoral dissertation, Illinois State University, 1965). Dissertation Abstracts, 1967, 27, 17034. (University Microfilms No. 66-5370)
- Gagné, R. M., & Briggs, L. J. Principles of instructional design. New York: Holt, Rinehart and Winston, 1974.
- Gerlach, V. S., & Ely, D. P. Teaching and media--a systematic approach (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall, Inc., in press.
- Irwin, J. L. An investigation of the effects of student self-evaluation and marking on achievement of cognitive and affective objectives of a basic mechanical drawing course. (Doctoral dissertation, Pennsylvania State University, 1973). Dissertation Abstracts International, 1973, 34, 1484A. (University Microfilms No. 73-24, 013)

- Kibler, R. J., Barker, L. L., & Miles, D. T. Behavioral objectives and instruction. Boston: Allyn and Bacon, Inc., 1970.
- Kibler, R. J., & Bassett, R. E. Writing performance objectives. In L. J. Briggs (ed.), Instructional design, principles and applications. Englewood Cliffs, N.J.: Educational Technology Publications, Inc., 1977.
- Kulhavy, R. W., Yekovich, F. R., & Dyer, J. W. Feedback and response confidence. Journal of Educational Psychology, 1976, 68, 522-528.
- Mager, R. F. Preparing instructional objectives. Palo Alto, CA: Fearon Publishers, 1961.
- McEwen, R. H. A comparison of four different techniques for improving performance through student self-evaluation (Doctoral dissertation, East Texas State University, 1967). Dissertation Abstracts, 1967, 27, 4049A. (University Microfilms No. 67-7356)

APPENDIX A

IME 411
 SS-1-78
 Planning Document Evaluation

NUM 1 2 3

NAME _____

Directions -- Using a 1 to 10 scale, with 1 being the lowest and 10 the highest, evaluate each section of the planning document. Please use only whole numbers (no fractions). Base your evaluations on the criteria supplied for each section. Each area has a blank to account for any other considerations which may come up. Thank you.

- (1-10) _____ 1 - Content Outline
- a) Is the content of your outline fully explained?
 - b) Is the skill or concept being taught covered in its entirety?
 - c)
- (1-10) _____ 2 - Need - Purpose
- a) Is the statement of the purpose a clear statement of the teacher's goal?
 - b) Is the need for the unit clearly described?
 - c)
- (1-10) _____ 3 - Audience Analysis
- a) Does the audience analysis contain some or all of the following:
 - 1) class size?
 - 2) description of ability level of students?
 - 3) description of racial makeup of the class?
 - 4) description of different socio-economic groups found in class?
 - 5)
- (1-10) _____ 4 - Objectives
- a) Are objectives meaningful?
 - b) Are objectives related to unit of instruction?
 - c) Are objectives:
 - 1) student oriented?
 - 2) behaviorally stated?
 - d) Do objectives:
 - 1) specify conditions of performance?
 - 2) specify standards of performance?
 - e)
- (1-10) _____ 5 - Media Selection
- a) Does media have potential for helping learners?
 - b) Are there other more appropriate media which could make the instruction more effective or efficient?
 - c) Are there any special circumstances that affect media selection which have not been considered?
 - d)
- (1-10) _____ 6 - Utilization
- a) Is the process for implementing the unit clear and sensible?
 - b) Are all the materials utilized?
 - c) Is the teaching strategy consistent with the audience, objectives, purpose, and time constraints?
 - d)

IME 411
SS-1-78
Materials Evaluation

NUM 1 2 3

NAME _____

Directions -- Using a 1 to 10 scale, with 1 being the lowest and 10 the highest, evaluate each aspect of the materials produced for the media product. Please use only whole numbers (no fractions). Base your evaluation on the criteria supplied for each aspect. Blanks are left to account for any other considerations which may come up. Thank you.

TECHNICAL QUALITIES CRITERIA

- (1-10) _____ 1. Lettering
- a) Is lettering legible?
 - b) Are letters neatly done?
 - c) Are letters and words appropriately spaced?
 - d) Are there any unwanted spots of ink or ink smears?
 - e) Have guidelines been completely erased?
 - f)
- (1-10) _____ 2. Illustration
- a) Is inking neat and accurate?
 - b) Is there sufficient contrast between original/board?
 - c) Have all pencil/stray marks been cleaned off?
 - d) Does enlargement fit board properly?
 - e)
- (1-10) _____ 3. Mounting
- a) Is visual completely adhered to backing?
 - b) Is visual trimmed neat and clean?
 - c) Have registration marks been erased?
 - d) Has excess MT-5, Sealamin, and rubber cement been removed?
 - e) Is visual well-positioned?
 - f)
- (1-10) _____ 4. Transparency making
- a) Are lines solid when projected?
 - b) Are colored areas solid and clean?
 - c) Is hinging secure and workable?
 - d) Is transparency neat and clean?
 - e)
- (1-10) _____ 5. Construction
- a) Is product constructed soundly or held together securely?
 - b) Has adequate caution been taken to assure product will stand up to classroom use?
 - c)
- (1-10) _____ 6. Appearance
- a) Is the product neat?
 - b) Have any errors been made which detract from the product's overall appearance?
 - c)
- (1-10) _____ 7. Other (please list)
- a)
 - b)

IME 411
 SS-1-78
 Mass Media Project Evaluation

NUM 1 2 3

NAME _____

Directions -- Using a 1 to 10 scale, with 1 being the lowest and 10 the highest, evaluate each aspect of the mass media product. Please use only whole numbers (no fractions). Base your evaluation on the criteria supplied for each aspect. Blanks are left to account for any other considerations which may come up. Thank you.

MASS MEDIA PROJECT CRITERIA

- (1-10) _____ 1. Need-Purpose
- a) Is the statement of the purpose a clear statement of the teacher's goal?
 - b) Is the need for the unit clearly described?
 - c)
- (1-10) _____ 2. Audience Analysis
- a) Does the audience analysis contain some or all of the following
 - 1) class size?
 - 2) description of ability level of students?
 - 3) description of racial makeup of the class?
 - 4) description of different socio-economic groups found in class?
 - 5)
- (1-10) _____ 3. Description of Program
- a) Title
 - b) Time
 - c) Where can it be viewed?
 - d) Length
 - e) Medium (TV, Film, etc.)
 - f)
- (1-10) _____ 4. Synopsis of Program
- a) Is program clearly described?
 - b) Is focus of lesson identified?
 - c)
- (1-10) _____ 5. Instructional Activities
- a) Do preprogram activities set up a framework for viewing?
 - b) Is the relationship between preprogram activities, postprogram activities and the program itself clear?
 - c) Do activities help learner attain the objective of the lesson?
 - d)
- (1-10) _____ 6. Objectives
- a) Are objectives meaningful?
 - b) Are objectives related to unit of instruction?
 - c) Are objectives:
 - 1) student oriented?
 - 2) behaviorally stated?
 - d) Do objectives:
 - 1) specify conditions of performance?
 - 2) specify standards of performance?